Ontology summarization: an analysis and an evaluation

Ontology Summarization: An Analysis and An Evaluation

Ning Li, Enrico Motta, Mathieu d'Aquin

Knowledge Media Institute
The Open University
Milton Keynes, United Kingdom
{n.li, e.motta, m.daquin}@open.ac.uk

Abstract. Ontology summarization has been recognized as a very useful technique to facilitate ontology understanding and then support ontology reuse as a new or supplementing technique. A number of efforts have emerged lately that apply different criteria, addressing different features of ontology, to extract ontology summaries. However, those efforts are ad-hoc in that there lacks consensus on a number of issues fundamental to the development of the field, such as a definition for ontology summarization, use case scenarios etc. Also, there lack sufficient evaluations and analysis, e.g. comparison among them and with other similar techniques, to provide meaning guidelines for users of this technique. With the aim to provide solutions to those fundamental issues, in this work, we present an analysis of this technique and its approaches. With the help of an objective evaluation method, we investigate what features of ontology are important in ontology summarization.

Keywords: Ontology, Ontology Summarization, Analysis, Evaluation

1. Introduction

The Semantic Web is growing fast and is rapidly emerging as a large-scale platform for publishing and sharing formalized knowledge models. Ontology understanding is important in ontology engineering to support tasks like ontology selection and reuse in constructing new ontology. This has been helped by the development of hierarchical-based ontology visualisation and navigation tools, such as OWLViz[1], OntoViz[2] and NeOn ontology visualiser[3]. However, with the size of ontology increasing as well as complexity of ontology taxonomy, not only representing ontology as tree elements were generally found to be a poor metaphor for user needs [1][2], also, the surveyed user experience on ontology engineering toolkits such as Protégé has found that such tools are too complex and do not reflect users' models of what they would expect to see in unfamiliar ontologies [3]. This becomes more problematic when users with limited ontology engineering experience encounter large

[1] http://www.co-ode.org/downloads/owlviz/
[2] http://protegewiki.stanford.edu/index.php/OntoViz
[3] http://www.neon-toolkit.org/wiki/1.x/OWL_Ontology_Visualization

ontologies in number, size as well as complexity. These observations are the motivations behind the work of developing novel interactive frameworks for ontology visualization and navigation [24] based on ontology summarization, which, in fact, has been recognised in recent years as an important tool to facilitate ontology understanding and help users quickly make sense of an ontology [4][5][6].

Apparently, ontology summarization shares a similar target with other ontology trimming/winnowing technologies, such as ontology partitioning [7][8], ontology modularization [9], ontology segmentation [10][11], application-driven ontology winnowing [12] etc., that is, to reduce the size and/or complexity of ontology to the level of necessary judged by either needs of users or requirements of tasks, and hence ease the burden of ontology management tasks. However, like all those technologies which approach the target from perspectives biased towards certain aspects of ontology, or geared towards applications/scenarios that rely on the techniques, ontology summarization, intuitive to its definition, has unique ways to approach the target and support applications/scenarios that depend on it.

While there is a clear need for ontology summarization, none of the work seen in literature has provided a well-defined meaning for it and thus differentiated it from other seemingly similar techniques, nor do they have a shared, but rather ad-hoc, understanding of what particular aspects of ontology are important or what determines the summary qualities etc. The lack of understanding on such fundamental issues undoubtedly hinders the development of the field. On one hand, it is difficult to appreciate the specialty of ontology summarization from other seemingly similar techniques. On the other hand, it is impossible to compare different approaches among them and provide users with a guideline of how to use this technique and its different approaches.

In this paper, we contribute to the development of ontology summarization from the following three aspects. Firstly, we take a step back from existing ad-hoc approaches to ontology summarization, provide a definition for it and clear ambiguities among different understandings with the help of exemplar use case scenarios. This is written in Section 2. Secondly, we come back to the state-of-the-art approaches on ontology summarization that aim to facilitate ontology understanding, to which we refer as user-driven ontology summarization, analyze them comparatively from the perspective of ontology features being addressed by those approaches. This is described in Section 3. Lastly, in Section 4, we design an evaluation process to find out which summarization criterion that features particular aspect(s) of ontology is more important than the others, and therefore, provide hints for practitioners about which approach to choose under what circumstances. This is followed by a discussion and conclusion of the paper in Section 5.

2. Ontology Summarization

2.1 Ontology Summarization Definitions

By the definition of "summary" in natural language processing given in [13], the features of summary include: 1) summaries may be produced from a single document or multiple documents; 2) summaries should preserve important information; 3) summaries should be short, no longer than half of the original text(s) and usually significantly less than that. In the context of ontology engineering, it is the second feature that fundamentally differentiates ontology summarization from other similar techniques. Though they also aim to reduce the size or complexity of original ontology significantly, instead of keeping "important" information, and more precisely "important for the whole ontology", they keep part or sub-topic information of ontology. For example, ontology partitioning and ontology modularization both concern the monolithic character of ontology that makes not only reasoning, but also modeling and visualization of large ontology extremely difficult [7]. Ontology partitioning approaches those problems by split one large ontology to many self-contained smaller sub-ontologies with each covering a certain subtopic, which, if put together again, form the original ontology to allow its easier maintenance and use [7][14], while ontology modularization focuses on selective use and re-use of smaller part of an ontology that covers certain aspects of the original ontology. Furthermore, ontology summarization should be "automatic", as text summarization [15], instead of semi-automatic relying on a trigger from a user or an application, which is often the case for other techniques [10][11][12]. Based on those features, we give a definition of ontology summarization, inspired by the text summarization definition in [15], as "the process of **automatically** creating a compressed version of a given ontology that provides **important** information for **the user**".

2.2 Scenarios for Ontology Summarization

A typical scenario in which a need for *ontology summaries* arises concerns ontology development, where a user may wish to use a semantic search engine, e.g., *Watson*[4] to locate and then explore ontologies which may provide conceptualizations relevant to the current model characterizing some particular entities. In such a scenario, a user can greatly benefit from *ontology summaries*, which, as the format to present the searching results to the user [16], helps him/her to quickly understand and compare candidate ontologies. This reinforces the point made by N. Noy that objective evaluations do not often support the ontology users to their best and that particular care should be taken to help naive users find ontologies and evaluate their suitability for the user's tasks [17]. A similar scenario where *ontology summaries* are very useful concerns online ontology sharing systems like Cupboard [18], which provides users

[4] http://watson.kmi.open.ac.uk

with their personal ontology spaces, where upload, share, review and connect ontologies are enabled. In such a scenario, snapshots of *ontology summaries* could provide a view to help user grasp what each ontology is about.

Also, *ontology summaries* have been used in an interactive ontology visualization and navigation tool, referred to as Key Concept Visualizer (*KC-Viz*)[5] using approaches in [5] with the details given in Section 3.1. A snapshot is presented in Fig. 1, where only *ontology summary*, in the form of ten key concepts, is shown for ontology *aktors portal*[6] containing hundreds of concepts. The size of the blue hexagon associated with specific key concepts is meant to represent the level of importance of the concepts. Each key concept is followed by a label containing its name and two numbers in brackets that represent the number of direct and indirect subclasses of the key concept. If users are interested in exploring the ontology further, they can extend/hide the visualization by integrating other entities related to the identified key concepts. A number of controls, which are self-explanatory in the figure, are provided to facilitate the visualization and navigation process.

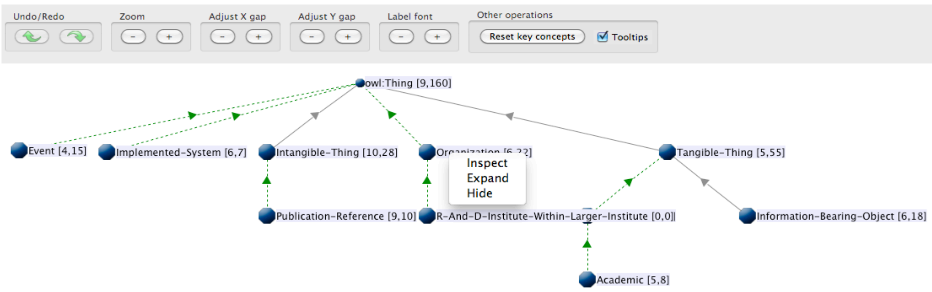

Fig. 1. Snapshot of Key Concept Visualizer (KC-Viz).

3. Ontology Summarization: An Analysis

Given that ontology entities and texts in natural language processing bear a similar feature of being either a collection of lexical labels or a bunch of sentences, a lot of experiences can be gained from text summarization to do ontology summarization. The first work, done by Zhang et al., looked into ontology summarization indeed from this perspective [4]. The authors, motivated by the work of a graph-based text summarization [15] and a semantic network analysis on ontology [19], take RDF sentence as the basic distilling unit for summarization and extract the most salient/important ones as summaries. This was followed by a second work which extracts only key concepts into summaries as better representatives of ontology [5]. In fact, concepts have been used as the representative entities by many ontology engineering tools, for example in semantic search engine swoogle[7], concepts are used to present the search results in a ranked order. Another work, also from Zhang et al.

[5] http://www.neon-toolkit.org/wiki/KC-Viz
[6] http://www.aktors.org/ontology/portal
[7] http://swoogle.umbc.edu

[6] extends the information used for the selection of most salient RDF sentence from those within a particular ontology to those harvested from Semantic Web. A feature in common, among these three approaches, and the only three to the best of our knowledge, is they all applied a number of criteria, with corresponding algorithms, either altogether with each algorithm addressing one particular feature of ontology as in [5], or separately with no clear indication of what features of ontology are particularly addressed by each algorithm as in [4][6]. Though the accumulated effect is subjectively evaluated as promising, that is, the algorithms-produced summaries approximate well to those manually selected by human assessors, there lacks an insightful view of what features of ontology play important role(s) in making some of entities into summaries while others not. In this paper, we apply an objective evaluation method to test against the key concept extraction approach where, as said, each algorithm addresses a particular feature of ontology [5]. Before this, we will provide some details of those algorithms and give our analysis of the features being considered in ontology summarization.

3.1. Approaches: a description

In [4], Zhang et al. took RDF sentences as the basic distilling unit for summarization. In other words, the summarization results are comprised of RDF sentences. By constructing an RDF sentence graph with RDF sentences as vertices and links among them as edges, the authors calculate, for each vertex, a "centrality" value that determines the relative importance of a vertex within the graph. The vertices, and thus the corresponding RDF sentences, with highest importance values will be extracted as ontology summaries. The "centrality" value of a sentence was determined using a number of criteria, which have been popularly used in the analysis of social networks. For example, *In-degree centrality* of a vertex measures the number of links to the vertex, which is generally interpreted as a form of *popularity* in social network and correspondingly, *out-degree centrality* would measure the number of links from the vertex to others, interpreted as *authority*. The link between two vertices *S1* and *S2* is established by the authors in the simplest term as follow: *if object of S1 is also subject of S2, then a link is established from S1 to S2*. *Betweenness centrality* of a vertex measures the occurrence of the vertex on the shortest paths between other vertices, that is to say, the more time a vertex occur in the shortest path between other vertices, the higher *betweenness centrality* value for the vertex than for others. In this particular context, RDF sentences with high *betweennness centrality* can be seen as "bridges" between clusters of RDF sentences. Thirdly, three other "centrality" measures, based on the *eigenvector* of the RDF graph, are used to provide more "centrality" values which address the structural and linguistic features of ontology. This approach, using RDF sentence as basic distilling unit instead of terms (i.e. concepts), the authors claim that it provides extra knowledge of how the terms are related in ontology and therefore provides a more comprehensive understanding of the ontology. However, as criticized in [16], when solely treating ontology as graphs and analyzing it with structural metric, the semantics down to concept level is ignored.

In a later work [6], Zhang et al., though still taking RDF sentences as the basic distilling unit, explored semantics of terms, e.g. subject, predicate or object, contained in RDF sentences to decide the salience of RDF sentences. It extended the RDF sentences in the ontology with “neighbouring information” by detecting how often the terms in the RDF sentences are linked or instantiated in global semantic web, that should illustrate the importance of the RDF sentences. However, the expansion was only made by a limited number of steps, three in this particular work and therefore is not really ‘global’ yet. Two “importance” measures are used to measure the salience of the RDF sentences in a global view. Firstly, the ***structural importance*** measures how many global semantic web entities have a *reference* to the local RDF sentences with regards to subjects, predicates or objects. Secondly, the ***pragmatics importance*** actually measures the statistics of terms being instantiated by other entities across global semantic web and thus indicates the *popularity* of terms appeared in local RDF sentences. This work, as declared by the authors, is more general and intuitive than the work in [4] because terms in RDF sentences can be used to influence the final results, whereas in [4], the RDF sentence is the smallest working unit.

The results in [4][6] are a certain number of the most salient/important RDF sentences in textual format. The fine-grained characters of ontology, say at concept level, have not been exploited to its full potential. In [5], ontology summarization was approached to extract key concepts by working on the primary entity of ontology, i.e. atomic classes (concepts) and the intrinsic relations among them. A number of criteria were jointly considered, and correspondingly a number of algorithms were developed and linearly combined, to identify key concepts of an ontology. Notably, the notion of ***natural category*** [20] was used to identify concepts that are information-rich in a psycho-linguistic sense. This notion was approximated by means of two operational measures: *name simplicity* which favors concepts that are labeled with simple names while penalizing compounds; and *basic level* which measures how ‘central’ a concept is in the taxonomy of the ontology. Two other criteria were drawn from the topology of an ontology: the notion of ***density*** highlights concepts which are information-rich in an ontological sense, i.e., they have been richly characterized with properties and taxonomic relationships while the notion of ***coverage*** aims to ensure that no important part of the ontology is neglected. Lastly, the notion of ***popularity***, drawn from lexical statistics, is introduced as a criterion to identify concepts that are commonly used in natural language. The key concepts were extracted depending on the final score of each concept which is a linear summation of the scores produced by each algorithm.

3.2 Approaches: An Analysis

The existing approaches to ontology summarization are ad-hoc in the sense that there lacks consensus on issues fundamental to the development of the field as a whole. First of all, the basic distilling unit of summarization is different. Secondly, different criteria, deemed to suit respective context most, are chosen for summarization, and therefore there lack foundations to compare those approaches. Thirdly and by no means the last, different names were given to the same criteria by different approaches which literally address the same feature of ontology, or the criteria of the same name, presumably address the same feature of ontology, were approached differently, evident from the *popularity* measure in approach [5] and [6]. Not only is this causing confusion to the users of this technique, also, it hinders the further development of the field. Here, we aim to provide a comprehensive view of ontology summarization from the following perspectives:

1. ***What features of ontology are being addressed***?

   We strongly believe that the main purpose of ontology summarization, unlike other ontology trimming techniques, is to facilitate users quickly make sense of ontology, meanwhile, using as few spaces as possible. Therefore, it is neither desirable nor necessary to keep complex, i.e. non-atomic entities, in summaries. This is especially important for none-experienced users of ontology. Therefore, we suggest **linguistic** aspects of ontology as the primary feature to be looked at in ontology summarization, such as name simplicity [5], term popularity in the scope of Web [5], which is a typical representation of natural language, or in the scope of Semantic Web [6], which is a typical metaphor, i.e., a formal explicit specification of a shared conceptualization, of natural language domain. This is also reflected in the **structural** aspects of ontology, such as hierarchy or taxonomy. If a concept is a hub connecting or a centre franchising many others, it is most probably that it is referred to by others more and hence more popular among others. There could be many other ways of using structural information. For example, *density* criterion looks into how a concept is richly described in terms of *is-a* and instantiation relations and *coverage* criterion makes sure maximum coverage of the ontology.

2. ***What criteria are being used***?

   Coherent with the ontology features being addressed, the criteria used to select summaries are tightly linked to those features. For example, *density* and *coverage* criteria [5], and *betweenness centrality* criterion [4] etc. are applied to the **structural** aspects of ontology while *name simplicity* [5] and *popularity* [6], *references* [5][6] etc. are applied to the **linguistic** aspects of ontology. Note that, since ontology summarization aims to find the "important" information for the whole ontology, some of the criteria used in ontology partitioning/modularization, such as covering sub-topic information, are not applicable.

3. ***How criteria are practiced?***

   Even the same criterion relating to the same feature of ontology is used, there could be more than one way of approaching it. Candidate approaches could vary in *the way how the algorithm is designed*, for example, the *popularity* can be calculated from information of Semantic Web [6], or non-Semantic Web [5], or *whether it relies on external knowledge*, that is knowledge harvested from Semantic Web or local to the ontology under question. For example, in [6], the authors rely on other ontologies collected from Semantic Web to decide the *reference* and *popularity* values of the terms in an RDF sentence. Also, in [5], the authors calculate the *popularity* value of each concept by counting the number of hits that returned when querying Yahoo with the name of the concept as keyword.

4. ***How the results are evaluated?***

   Just as experiences can be gained from text summarization to do ontology summarization, lessons can be learnt from the evaluation of text summarization to evaluate ontology summarization. Also, as it is ontology summarization, some of the evaluation techniques for ontology are applicable to ontology summaries. This was investigated and a comparative evaluation among the approaches is given in [21]. Before that, the evaluations were undertaken in an ad-hoc manner.

By now, a systematic view of ontology summarization technique and its approaches have been given. We will then focus on, by means of an evaluation, investigating the impact of the features, embodied into criteria, on the summarization results, i.e. summaries, with respect to each other. As been emphasized throughout this paper, with the final summarization result being an accumulated effect of a series of criteria encapsulating different features of ontology, as seen in approaches [5] and [6], it is not possible to separate the impact of each criterion, and thus each feature of ontology, on making results a good summary, which is judged by comparing it with the one manually selected by human assessors. Hence, there is a need to split the criteria, comparatively evaluate them and find out what features of an ontology make some entities into summaries while leaving others out. This will be described next.

4. Impact of Ontology Features: An Evaluation

4.1 Evaluation settings

The setting of our evaluation is as follows: eight people, each with good experience on ontology engineering, were asked, for each ontology, to manually extract up to 20 key concepts they considered the most representative for summarizing the contents of the ontology. The concepts that were chosen by at least 50% of the experts form a reference summary, referred to as "ground truth" summary. This will be used later in the analysis of evaluation results. Two ontologies, *biosphere*[8], *financial*[9] were used, which have also been used in [5][21] and contain no properties or instances, and thus provide a rather clean environment because we summarize concepts only.

We use two criteria ***density*** and ***reference*** as embodiments of the **structural** features of ontology and another two criteria ***popularity*** and ***name simplicity*** reflecting the **linguistic** features of ontology. We then run through an evaluation process to find out the order of importance of these criteria, which provides answers to the most important question this paper aims to answer, that is, what features of ontology are thought important in ontology summarization. First of all, we introduce the implementations of criteria involved in the evaluation one by one.

***Density*:** The $density(C) \in [0..1]$ of a concept C is a measure of how richly described the concept is in ontology and is computed on the basis of its number of direct sub- concepts, properties and instances. In the context of this evaluation, it counts the number of *is-a* relations on concepts only.

***Reference*:** The $reference(C) \in [0..1]$ of a concept C provides a normalized measure of the number of entities dynamically collected from Semantic Web using semantic search engine Watson, which reference (depend on) the concept C. It counts the axioms which have the concept on the right-hand side, i.e., the number of assertion <*s, p, o*> such that o is the considered concept C. Those axioms potentially involve property *domain* and *range* as well as *instantiation* relations besides the *is-a* relations because ontologies collected from Semantic Web may contain those relations, though

[8] http://sweet.jpl.nasa.gov/ontology/bioshpere.owl
[9] http://www.larflast.bas.bg/ontology

our experimental ontologies do not. Therefore, ***reference*** should provide a more precise indication of how dense a concept is described in the scope of Semantic Web.

***Name simplicity*:** The *name simplicity*, $NS(C) \in [0..1]$ is 1 if the label of concept C is made of only one word. It decreases following the number of compounds in the label, in accordance with the following formula: $NS(C) = 1 - c(nc-1)$, nc being the number of compounds in the label and c a constant in our experiments, we use $c = 0.3$. For example, the name simplicity of the concept *Artist* is 1, while that of *MusicalArtist* is 0.7. The rationale for this criterion is that natural categories normally have relatively simple labels, such as chair or cat. That is, they are unlikely to be compound terms.

***Popularity*:** The *popularity* $(C) \in [0..1]$ is a normalized number of results returned by querying Yahoo with the name of C as keyword. Compound names are transformed to a sequence of keywords separated by a space. The rationale behind this criterion is that concepts generally share the same meaning as they are in natural language and we should try to identify concepts that are particular common in natural language.

4.2 Evaluation

Kendall's tau [22] (abbr. as *tau*) coefficient is often used to measure the agreements between two measured quantities. In specific, it is a measure of rank correlation, that is, the similarity of the orderings of the data when ranked by each of the quantities. It has been used as sentence-rank-based evaluation tools for text summarization [23] as well as ontology summarization [4][21]. Here, we use *tau* to find the correlation between the score vector (one per ontology and the length of vector equals the number of concepts in each ontology), produced by each criterion, with "ground truth" score vector. The score vector for each criterion is obtained by running the corresponding algorithm. Different from the "ground truth" summary, the "ground truth" score vector is obtained by counting the eight experts' votes on each concept and then normalizing the result with respect to the total number of votes being cast to the whole ontology. In this case, when a concept receives no votes, its score value in the "ground truth" is zero. We evaluate the criteria described in above section. Table 1 shows the *tau* scores, where each entry is a *tau* score indicating the rank correlation between the corresponding criterion score vector and "ground truth" score vector.

Table 1. Agreement between each criterion and ground truth measured by *tau*

	Density	**Name Simplicity**	**Popularity**	**Reference**
Biosphere	**0.454**	0.111	0.3	**0.456**
Financial	**0.539**	0.464	0.43	**0.517**
Average	**0.497**	0.288	0.365	**0.487**

Note that the resulted *tau* score does not reflect the precise importance, rather a rank of importance, of each criterion in making the algorithm results close to "ground truth" [4]. Increasing values imply increasing agreement between the two sets of rankings, i.e. algorithm results ranking and "ground truth" ranking. In the case that the rankings are completely independent and uncorrelated, the coefficient will have value zero on average. Apparently, if one criterion consistently produces higher scores than other criteria cross all ontologeis, it is reasonable to believe that it is a more important criterion and would have a higher average *tau* score. The average score of each criterion over the two ontologies is listed in the bottom row of Table 1.

From the results, we can see that, the ***density*** and ***reference*** criteria rank among the highest for both ontologies with ***density*** being marginally higher than ***reference*** in average. This is a very interesting finding. It shows that, even if ***density*** uses only the *is-a* relations local to the ontology while ***reference*** uses all the relations, collected from Semantic Web by Watson semantic search engine, where the concept under scrutiny appears as an object, the summary produced using the criteria ***density*** ends up with a higher average *tau* agreement score with "ground truth" than the criterion ***reference*** does. This is not surprising because we are measuring the agreement with "ground truth" that is produced by human assessors who only have the knowledge of local ontology. The ***reference*** criterion as a measurement of ***density*** in a global sense is not rightly reflected here. This may highlight the limitation of subjective evaluation approaches which rely on subjective opinions and have been popularly used in many areas including text summary evaluation and ontology evaluation [21].

The order of rankings between the remaining two criteria varies across the two ontologies. Though the average score at the bottom row provides a more comprehensive indication of the importance of each criterion, a closer look into those variations could provide a profound insight into the impact of the criterion on ontologies with distinctive features. For example, the ranking of *name simplicity* is lower than *popularity* in *biosphere* ontology but higher in *financial* ontology. So, why, in another word, *name simplicity* is less important than *popularity* in *biosphere* ontology but more important in *financial* ontology. Firstly, let's look at what's typically contained in *biosphere* ontology as illustrated in Fig. 2 using *KC-Viz*. A majority of the terms are simple names instead of compounds. Furthermore, a high percentage of the terms are not very commonly used, and therefore would have a low *popularity* value. The popular terms mostly appear in a place which could end it with a high '*density*' value, as seen in Fig. 2. Therefore, the impact of *name simplicity* is less prominent than that of *popularity* in making the summarization results correlating with "ground truth" summary, which contains ten key concepts, i.e. *Animal, Bird, Fungi, Insect, Mammal, MarineAnimal, Microbiota, Plant, Reptile, Vegetation*, all with very popular names and only one is compound.

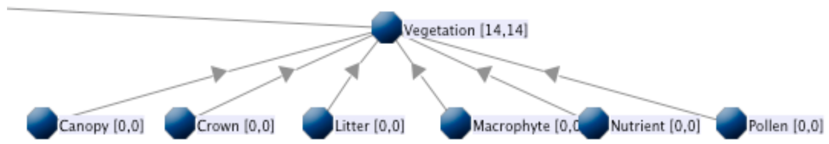

Fig. 2. A snapshot of *biosphere* ontology.

For *financial* ontology, a majority of the terms are labeled with popular words whose popularity values differ less significantly than those in *biosphere* ontology. It is often the case that a simple name is franchised by many compound names, as shown in Fig. 3. With nine key concepts in "ground truth" summary containing only one compound name, i.e. *Bank, Bond, Broker, Capital, Contract, Dealer, Financial_Market, Order, Stock*, it is not surprising that *name simplicity* impose a larger impact than *popularity* on the results in making them correlate with "ground truth" summary more.

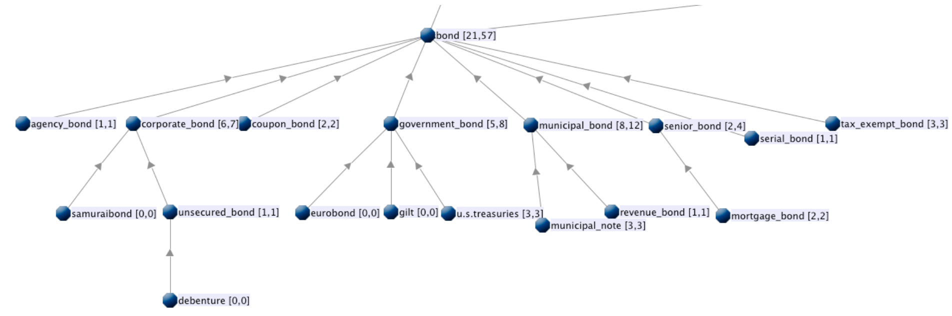

Fig. 3. A snapshot of *financial* ontology.

5. Discussions and Conclusions

This paper firstly addressed the fundamental issues in the field of ontology summarization, which have been overlooked by literatures. That is, to identify the purpose and use case scenarios of ontology summarization; provide a definition for it; identify the special characters which differentiate it from other seemingly similar techniques. By analyzing the state-of-the-art approaches, we provide a comprehensive view of this technique from a number of perspectives. We then focus on the investigation of what particular features of ontology are important and should be considered in ontology summarization, and how to approach them, what determines the summary qualities etc. An evaluation is designed to find the impact of using different criteria that address different features of ontology. The evaluation given focused on the extraction of key concepts using two ontologies which contain only concepts. It could be extended to include key properties or key instances into summaries if a use case scenario, such as driven by applications, is envisioned. In the context of user-driven ontology summarization whose primary target is to facilitate users ontology understanding, such an extension is not seen as a requirement.

A crucial issue that remains controversial and will certainly drive future research on ontology summarization is evaluation, as happened in text summarization domain. The creation of training material sets and the establishment of baselines for performance levels are challenging and remain absent. More collaborative research efforts are required.

References

1. Dzbor, M., Peroni, S., Motta, E., d'Aquin, M.: NeOn Toolkit Plug-in for Visualization and Navigation in Ontologies and Ontology Networks Based on Concept Summarization and Categorizing. NeOn Project Deliverable D4.5.4, Feb. (2009)
2. Duineveld, A.J. et al.: WonderTools? A Comparative Study of Ontological Engineering Tools. Intl. J. of Human-Computer Studies. **52**(6), pp.1111-1133 (2000)
3. Storey, M.A., Lintern, R., Ernst, N.A., et al.: Visualization and Protégé. In: 7th International Protégé Conference. Maryland, US (2004)

4. Zhang, X., Cheng, G., Qu, Y.: Ontology Summarization Based on RDF Sentence Graph. In: 16th International World Wide Web Conference, Banff, Alberta, Canada, May 8-12 (2007)
5. Peroni, S., Motta, E., d'Aquin, M.: Identifying Key Concepts in an Ontology Through the Integration of Cognitive Principles with Statistical and Topological Measures. In: 3rd Asian Semantic Web Conference, Bangkok, Thailand (2008)
6. Zhang, X., Cheng, G., Ge, W., Qu, Y.: Summarizing Vocabularies in the Global Semantic Web. Journal of Computer Science and Technology 24(1), pp.165-174. Jan. (2009)
7. Stuckenschmidt, H., Klein, M.: Structure-based Partitioning of Large Concept Hierarchies. In: 3rd Int. Semantic Web Conf. (ISWC), Hiroshima, Japan. (2004)
8. MacCartney, B., McIlraith, S., Amir, E., Uribe, T.E.: Practical Partition-Based Theorem Proving for Large Knowledge Bases. In: Proc. of the International Joint Conference on Artificial Intelligence (IJCAI) (2003)
9. d'Aquin, M., Sabou, M., Motta, E.: Modularization: a Key for the Dynamic Selection of Relevant Knowledge Components. In: Workshop on Modular Ontologies, ISWC. (2006)
10. Seidenberg, J., Rector, A.: Techniques for Segmenting Large Description Logic Ontologies. In: Workshop on Ontology Management: Searching, Selection, Ranking, and Segmentation. 3rd Int. Conf. Knowledge Capture (K-Cap), pp. 49–56, Canada (2005)
11. Bhatt, M., Wouters, C., Flahive, A., Rahayu, W., Taniar., D.: Semantic Completeness in Sub-ontology Extraction using Distributed Methods. In: Proc. Int. Conf. on Computational Science and its Applications (ICCSA), pp. 508–517, Perugia, Italy (2004)
12. Alani, H., Harris, S., O'Neil, B.: Winnowing Ontologies Based on Application Use. In: 3rd European Semantic Web Conference (ESWC), Budva, Montenegro, June (2006)
13. Das, D., Martins, A. F.T.: A Survey on Automatic Text Summarization. In: Literature Survey for the Language and Statistics II course at CMU (2007)
14. d'Aquin, M., Schlicht, A., Stuckenschmidt, H., Sabou, M.: Criteria and Evaluation for Ontology Modularization Technique Criteria and Evaluation for Ontology Modularization Technique. In: eds. Stuckenschmidt, H., Parent, C., Spaccapietra, S., Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization (2009)
15. Erkan, G., Radev, D. R.: LexRank: Graph-Based Lexical Centrality as Salience in Text Summarization. Journal of Artificial Intelligence Research, 22, pp. 457–479 (2004)
16. Sabou, M., Lopez, V., Motta, E., Uren, V.: Ontology Selection: Ontology Evaluation on the Real Semantic Web. In: Workshop: Evaluation of Ontologies for the Web (EON) at 15th International World Wide Web Conference, Edinburgh (2006)
17. Noy, N. F.: Evaluation by Ontology Consumers. IEEE Intelligent Systems, 19(4):74 – 81, July/August (2004)
18. d'Aquin, M., Lewen, H.: Cupboard -- A Place to Expose your Ontologies to Applications and the Community. Demo at the 2009 European Semantic Web Conference (2009)
19. Hoser, B., Hotho, A., Jaschke, R., Schmitz, C., Stumme, G.: Semantic Network Analysis of Ontologies. In: Proc. of the 3rd European Semantic Web Conference, pp. 514–529 (2006)
20. Rosch, E.: Principles of Categorization, Cognition and Categorization. Lawrence Erlbaum, Hillsdale, New Jersey (1978)
21. Li, N., Motta, E.: Evaluations of User-driven Ontology Summarization. In: Proc. of 17th International Conference on Knowledge Engineering and Knowledge Management by the Masses (to appear) (2010)
22. Sheskin, D.J.: Handbook of Parametric and Nonparametric Statistical Procedures. CRC Press (1997)
23. Donaway, R.L., Drummey, K.W., Mather, L.A.: A Comparison of Rankings Produced by Summarization Evaluation Measures. In: ANLP/NAACL Workshop on Automatic Summarization, pp 69–78 (2000)
24. Peroni, S., Motta, E., Li, N., and d'Aquin, M.: Revision and Integration of Neon Toolkit Plug-ins for Visualisation and Navigation into NeOn infrastructure. NeOn Project Deliverable D4.5.5 (2010)